DOCUMENT RESUME

ED 185 759 EC 123 914

AUTHOR	McGue, Matthew: And Others
TITLE	Validity of the Woodcock-Johnson Psycho-Educational Battery with Learning Disabled Students.
INSTITUTION	Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
REPORT NO	IRLD-RR-15
PUB DATE	Nov 79
CONTRACT	300-77-0491
NOTE	56p.: See also EC 123 901-925.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Elementary Education: Exceptional Child Research: *Learning Disabilities: *Standardized Tests: *Testing Problems: *Test Validity
IDENTIFIERS	*Woodcock Johnson Psycho Educational Battery

ABSTRACT

The validity of the Woodcock-Johnson Psycho-Educational Battery was examined using test results of 50 learning disabled fourth graders. The appropriateness of the developmental strategy and the evidence for the external validity of the cluster measures contained in the battery were considered. Results indicated that the factor and scholastic aptitude measures had little external validity when using other standardized measures as criteria. It was suggested that the lack of demonstrable validity resulted from the questionable nature of the operational definitions of the test's constructs. (Author/CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED185759

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

University of Minnesota

Research Report No. 15

VALIDITY OF THE WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY WITH LEARNING DISABLED STUDENTS

Matthew McGue, Mark Shinn, James Ysseldyke

iRLD

Institute for Research on Learning Disabilities

EC143914

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

James Ysseldyke

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IRLD

Director: James E. Ysseldyke

Associate Director: Phyllis K. Mirkin

The Institute for Research on Learning Disabilities is supported by a contract (300-77-0491) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U.S. Office of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled children. Research activities are organized into eight major areas:

I. Adequacy of Norm-Referenced Data for Prediction of Success

II. Computer Simulation Research on the Assessment/Decision-making/Intervention Process

III. Comparative Research on Children Labeled LD and Children Failing Academically but not Labeled LD

IV. Surveys on In-the-Field Assessment, Decision Making, and Intervention

V. Ethological Research on Placement Team Decision Making

VI. Bias Following Assessment

VII. Reliability and Validity of Formative Evaluation Procedures

VIII. Data-Utilization Systems in Instructional Programming

Additional information on these research areas may be obtained by writing to the Editor at the Institute.

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Research Report No. 15

VALIDITY OF THE WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY WITH LEARNING DISABLED STUDENTS

Matthew McGue, Mark Shinn, James Ysseldyke

Institute for Research on Learning Disabilities

University of Minnesota

November, 1979

FEB 1 3 1980

Abstract

A critical examination of the validity of the Woodcock-Johnson Psycho-Educational Battery for the assessment of learning disabled children was made. This evaluation involved both a consideration of the appropriateness of the developmental strategy and the evidence for the external validity of the cluster measures contained in the battery. Using the test results of fifty learning disabled children, the factor and scholastic aptitude measures were found to have little external validity when using other standardized measures as criteria. This lack of demonstrable validity is seen as a result of the questionable nature of the operational definitions of the constructs involved. Specific recommendations were made for how the Woodcock-Johnson can be validly used in the assessment of learning disabled children.

Table of Contents

	Page
Introduction	1
Development of the Woodcock-Johnson Psycho-Educational Battery	2
Development of Clusters	3
Method	4
Research Questions	4
Validity of the Cognitive Factor Clusters	5
Validity of the Scholastic Aptitude Clusters	5
Validity of the Achievement Clusters	5
Subjects	5
Data Collection Procedures	6
Results	6
Cognitive Factor Clusters	6
Development	6
Validity of the Factorially-Derived Cognitive Clusters for LD Students	9
Subtest Composition and Weighting	11
External Validity	15
Scholastic Aptitude Clusters	17
Development	17
Subtest Composition and Weighting	19
External Validity	21
Achievement Clusters	22
Development	22
External Validity	24
Conclusions	25
References	28
Footnote	30
Tables	31

Validity of the Woodcock-Johnson Psycho-Educational Battery with Learning Disabled Students

The identification, assessment, and treatment of students who could benefit from special educational services has always been a major concern of the educational psychologist. Recently, this concern has been intensified through the legislative mandate provided by Congress (Education of all Handicapped Children Act, Public Law 94-142, 1975) which, among other things, directs that the learning disabled student be identified and validly assessed. Unfortunately, concerns for those who need special educational services exceed the availability of psychometrically adequate tests that can be used to identify learning disabled students (Ysseldyke, 1977). In particular, Ysseldyke argues that presently used instruments lack appropriate standardization samples, adequate reliability, and demonstrated validity when used with special populations. Furthermore, for the frequently used diagnostic sign of a discrepancy between aptitude and achievement, it is desirable that both the aptitude and achievement scores be normed on the same standardization sample, a condition Ysseldyke found was not often met in practice.

The recent publication of the Woodcock- Johnson Psycho-Educational Battery (W-J), a collection of wide-age-range, individually administered measures, appears to provide an instrument which meets many of the assessment requirements of educators and researchers who work with the learning disabled (Woodcock, 1978). The battery's aptitude and achievement measures were standardized on the same carefully-selected standardization sample, reliability evidence was extensively documented, and all aspects of validity were carefully considered. Furthermore, special care

was taken to ensure that the battery be appropriate for the assessment of learning disabled individuals. This concern is evidenced not only by the author's attempt to obtain validity information for the use of the battery with the trainable mentally retarded and learning disabled, but also in a consideration of these special populations in the original scale development.

Given the legislative directive, the paucity of technically adequate assessment instruments, and the careful developmental approach taken by Woodcock and his colleagues, it is little wonder that the W-J has been warmly received by researchers and educators alike.

The present paper evaluates the extent to which Woodcock has been successful in producing an instrument that can be used to validly assess learning disabled students. Although we appreciate the special difficulties involved in attempting to develop assessment devices which can be used in both special and normal populations, we believe that careful scrutiny of the W-J is warranted not only due to its popularity, but also due to the specific claims made by the author that the battery is appropriate for the identification and assessment of students with learning deficits and/or special problems or disabilities (Woodcock, 1978, p. 4).

Development of the Woodcock-Johnson Psycho-Educational Battery

In its entirety, the Woodcock-Johnson Psycho-Educational Battery contains 27 subtests; 12 measuring cognitive ability, 10 measuring scholastic achievement, and five designed to assess an individual's pattern of interests. In the present paper, we are concerned with 22 of the 27 subtests -- those used in the Cognitive Ability and

Achievement batteries. Table 1 lists these 22 subtests along with brief descriptions of the tasks involved in each. The 12 Cognitive Ability subtests represent a broad sampling of the cognitive domain, while the 10 Achievement subtests reflect those standard dimensions upon which our nation's educational curriculum is based.

- - - - - - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - - - - - -

To prevent generalizations from specific subtests, the subtests of the W-J are organized into clusters (weighted composites of two or more subtests) that are considered indices of more general functions. The 12 Cognitive Ability subtests are organized into eight Cognitive Ability clusters; the 10 Achievement subtests are organized into five Achievement clusters.

Consistent with Woodcock's position that interpretations should be based on broad rather than narrow measures, the norms, reliability, and validity of the W-J are reported for clusters rather than subtests.

Development of Clusters

Cluster development followed a careful sequential process. First, the construct to be measured by each cluster was identified. Second, those subtests which measured each construct were selected. Finally, the relative weighting of the subtests in forming the composite cluster was determined. The empirical procedures involved in cluster development were based upon the subtest results of some 2,000 individuals in the calibration-norming sample.

In developing the clusters, Woodcock employed three separate

strategies which resulted in qualitatively distinct measures. With the first strategy, Woodcock sought to identify and measure those dimensions underlying performance on the 12 Cognitive Ability subtests. These dimensions were identified by means of factor analysis, and consequently led to the definition of four factor clusters: Verbal Ability, Reasoning, Perceptual Speed, and Memory. Woodcock next sought to identify those Cognitive subtests which could be used to assess the potential for future achievement. Here, a multiple regression based strategy was used to develop four Scholastic Aptitude Clusters: Mathematics, Reading, Written Language, and Knowledge. Finally, the 10 Achievement subtests were rationally organized into five broad Achievement clusters: Reading, Mathematics, Written Language, Knowledge, and Skills. A listing of the 12 clusters along with their subtest composition and weighting is presented in Table 2.

- - - - - - - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - - - - - - -

Method

Research Questions

Because clusters provide the interpretative bases of the W-J, our evaluation of the battery examined the extent to which the strategies of cluster development resulted in measures that are both appropriate and valid for assessing learning disabled students. For each strategy of cluster development, we will present a more detailed description of the developmental procedures involved, and then a general discussion of the validity of the procedures. Finally, the following research questions

will be addressed:

Validity of the Cognitive Factor Clusters.

1. To what extent are the four factors identified by Woodcock the salient dimensions that underlie performance of learning disabled students on the 12 Cognitive Ability subtests (i.e., is the performance of LD students best represented by the four factors of verbal ability, reasoning, perceptual speed, and memory)?
2. To what extent is the adopted weighting system appropriate for assessment of learning disabled students?
3. To what extent are the Woodcock-Johnson factorially-derived Cognitive Ability clusters externally valid for learning disabled students?

Validity of the Scholastic Aptitude Clusters.

1. To what extent is the weighting system used in deriving the four Scholastic Aptitude clusters appropriate in assessing learning disabled students?
2. To what extent are the Scholastic Aptitude clusters externally valid?

Validity of the Achievement Clusters.

To what extent are the rationally-derived Achievement clusters valid for the assessment of the achievement of learning disabled children?

Subjects

Subjects were 39 male and 11 female fourth graders from metropolitan Minneapolis and St. Paul schools. All subjects were identified as

learning disabled by placement teams in the school districts they attended, and all assessment took place within six months of the initial identification. The average age of the subjects was 10 years, 1 month (SD = 5.04 months).

Data Collection Procedures

Pupils were assessed as part of a larger project comparing the test performance of learning disabled students to that of a group of 49 underachieving non-learning disabled students (Ysseldyke, Algozzine, Shinn, & McGue, 1979). Data from other standardized tests in the domains of learning aptitude (Wechsler Intelligence Scale for Children--Revised, [WISC-R], achievement (Peabody Individual Achievement Test [PIAT], Stanford Achievement Test--mathematics computation and mathematics concepts subtests), and perceptual-motor functioning (Bender Visual Motor Gestalt Test [BVMGT], Developmental Test of Visual-Motor Integration [VMI]) were collected concurrently. Students were tested in either three or four 90-minute sessions.

Results

Cognitive Factor Clusters

Development. In deriving the factor clusters, Woodcock sought to identify those dimensions, common across age levels, underlying performance on the 12 Cognitive Ability subtests. In order to achieve this goal, a factor analysis of the 12 subtests was performed at each of nine age levels (ranging from preschool through adult). Interpretation of the factor analytic results led to the identification of four factors relevant throughout the age span--verbal ability, reasoning,

perceptual speed, and memory. In order to develop the factor clusters, factor scores were estimated for each of the four factors at each age level for the 2,000 individuals in the developmental sample. These estimated factor scores were then used as dependent variables in subsequent analyses. Using half the developmental sample, Woodcock first determined cluster composition by selecting those Cognitive Ability subtests (between two and four) which best predicted these estimated factor scores in a stepwise multiple regression analysis. After cluster composition was decided, the test results on the remaining 1,000 individuals were used to develop the relative weighting of the subtests in the computation of the cluster scores. These weights were determined through a series of multiple regressions at each of the nine age levels with the selected Cognitive Ability subtests as the independent variables and the relevant estimated factor scores as the dependent variable. Finally, a common set of weights which reflected the results at each of the nine age levels was determined.

Woodcock's method for deciding the relative weighting of the subtests in the factor clusters is an attempt to derive an efficient and stable method of factor score estimation. Unfortunately, the test manual provides little information which would allow for an evaluation of the extent to which the proposed measures are valid for the constructs (factors) in question. For example, although we are told that the Numbers Reversed and Memory for Sentences subtests are the best predictors of the original factor score estimates, we have no ideas as to how well this composite does predict these estimates. Furthermore, we have no idea of how well the composite of Memory for Sentences and Numbers Reversed

predict the criterion of importance, the memory factor scores. If the linear composite of Numbers Reversed and Memory for Sentences had only a moderate correlation with the actual memory factor, then we could have little confidence in interpreting this measure in terms of this originally identified memory factor (see for example, Green, 1976).

Although this does not seem to represent a special problem when interpreting the Memory cluster score, as both subtests obviously involve a memory operation, it is somewhat problematic when considering the unique subtest composition and pattern of positive and negative weights that characterize the Reasoning and Verbal Ability Clusters. For example, one cannot through simple inspection of the subtest composition of the Reasoning factor cluster conclude that the positive weighting of the Analysis-Synthesis, Concept Formation, and Analogies subtests along with the negative weighting of the Antonyms-Synonyms subtest results in a measure of reasoning. Instead the interpretation of the factor cluster score in terms of the reasoning construct requires that: (a) the initially identified factor was indeed a reasoning factor, and (b) the linear composite of the subtests is such an accurate predictor of the actual factor scores that one would be willing to interpret the predicted score as if it were the actual criterion score. To illustrate the difficulties which might arise if the second condition is not met, consider the prediction of achievement test scores from intelligence test scores. Although we are able to predict performance on standardized achievement tests with a fair degree of accuracy using an individual's intelligence test score, we would not necessarily want to conclude that a change in measured intelligence, say due to some intervention, resulted

in a change in actual achievement. We would of course be justified in concluding that a change in measured intelligence results in a change in predicted achievement, but this would be of little help if the research question at hand involved a determination of the effect of intervention upon actual achievement. Thus, a general difficulty in using W-J factor clusters as dependent variables in any research endeavor is the restrictive nature of having to interpret one's results in terms of an effect upon predicted factor scores rather than actual factor scores. This difficulty is further confounded by the lack of information concerning the accuracy of this prediction.

Validity of the Factorially-Derived Cognitive Clusters for LD Students. The first research question addressed the extent to which the performance of learning disabled students on the 12 Cognitive Ability subtests was best represented by the factorially derived clusters of verbal ability, reasoning, perceptual speed, and memory. We examined the factor structure of the 12 Cognitive Ability subtests for our sample of 50 learning disabled students. Principal axes extraction (with iterated communality estimates) followed with a varimax rotation of the factor loadings was performed on the 12 subtests with either one, two, three, or four factors extracted. The four-factor solution did not yield the four factors identified by Woodcock. Since the varimax rotation produced an uninterpretable set of factor loadings, an alternative solution was sought. On the basis of both objective criterion (eigenvalues greater than one) and subjective criteria (interpretability of the results), the three-factor solution, which accounted for 45 percent of the variance, was indicated. Table 3 lists the varimax rotated factor loadings along with the final

communality estimates for each of the 12 subtests for the preferred solution.

- - - - - - - - - - - - - - - - - -

Insert Table 3 about here

- - - - - - - - - - - - - - - - - -

The first factor was obviously a verbal factor, as it loaded highly on all subtests with a verbal content. Although this first factor did coincide with one of the four factors found in the original development of the battery, the other two factors did not. The second factor had high or moderate loadings on all subtests requiring verification against some standard, whether that standard was a geometric form, as in the Spatial Relations subtest, or a reference key, as in the Analysis-Synthesis subtest. The final factor appeared to be defined in terms of content, since it had high to moderate loadings on all subtests requiring the manipulation of numbers. Interestingly enough, it did not seem to matter whether the manipulation of numbers was in a perceptual speed task such as Visual Matching, a memory task, such as Numbers Reversed, or in actual arithmetic operation, such as Quantitative Concepts.

Although we are well aware that the results presented above could be a function of sampling fluctuation, due to the moderate size of the present sample, we do believe that our findings suggest that an alternative structure should be entertained. That is, the population of learning disabled children might be characterized by a factor structure which differs from the factor structure evidenced in a large sample of normal individuals. In fact, the literature is replete with investigations demonstrating that special populations have discrepant

factor structures (e.g., Garai & Scheinfeld, 1968). It is not that those factors identified in a normal population are unimportant when one looks at special populations, but rather that these factors are evidenced in a different manner and to a different degree in any given set of variables. In short, a measurement technique which requires the differential weighting of factorially complex variables is apt to produce measures which have a low degree of generalizability. If the desire is to develop factor measures, it would be better to first identify the relevant constructs using factorially complex variables, and then develop a pure measure of the factor through an extensive content analysis of the factor analytic results. Such a strategy has been successfully employed by ETS in their development of their factor-referenced tests (Ekstrom, French, Harman, & Dermen, 1976). In the absence of a pure measure of the factor, one should then realize that those characteristics of their special population which they desire to determine, might precisely be those characteristics which make the use of the W-J factor clusters inappropriate.

Subtest Composition and Weighting. Even should one accept the factor analytic results, the method whereby subtest composition was decided appears problematic. The stepwise multiple regression criteria adopted by Woodcock resulted in factor clusters with a marked degree of overlap. Not only is that overlap evidenced by the moderate size of the intercorrelations between the factor clusters (see Table 4), but also by the actual subtest overlap between the clusters (see Table 2). For example, the Antonyms-Synonyms subtest is used in the computation of both the Verbal Ability cluster score and the Reasoning cluster score. In

fact, the Antonyms-Synonyms subtest is used in the computation of six of the eight Cognitive Ability scores, Analysis Synthesis in three, etc. Remarkably, not one of the eight Cognitive Ability clusters is composed entirely of subtests unique to that cluster. This overlap is a direct result of the test developers' emphasis upon the development of measures having high degrees of convergent validity, to the neglect of the measures' discriminant validity (Campbell & Fiske, 1959). For example, in selecting subtests to be used in measuring an individual's verbal ability, there is no preference for those subtests which measure verbal ability but not reasoning. Thus, to a certain extent the Verbal Ability cluster measures an individual's ability to reason. This overlap is a special source of difficulty for educational assessment in general, and the assessment of learning disabled children in particular. This is due to the emphasis in educational assessment placed upon a consideration of both the intellectual strengths and weaknesses of a child as indicated by observed differences in measured abilities. But, the cluster overlap in the W-J could serve to mask the existence of actual discrepancies. And even should a discrepancy be observed, caution is warranted due to the highly unreliable nature of differences between correlated measures (Harris, 1963). Finally, and perhaps most discomforting, those who would use the W-J for applied assessment should be aware that a specific cognitive deficit on, say, the Antonyms-Synonyms subtest, might result in a profile which suggests a rather general cognitive deficit.

- - - - - - - - - - - - - - - - - -

Insert Table 4 about here

- - - - - - - - - - - - - - - - - -

In addressing the question of the appropriateness of the weighting

system employed we are limited, by our inability to recover the original factor structure, to a discussion of general principles rather than specific evaluations. On the whole, the test developers take great care in insuring that their weights not be overly sample specific, using the parallel results at nine separate age levels to determine the common set of weights. But, one characteristic of the weighting scheme adopted does warrant special consideration: the use of suppressor variables in the computation of both the Verbal Ability and Reasoning cluster scores. To illustrate what is meant by a suppressor variable, consider the Reasoning cluster. In this cluster the Analysis-Synthesis, Concept Formation, and Analogies subtests all are weighted positively, while the Antonyms-Synonyms subtest has a negative weight. Although performance on the Antonyms-Synonyms subtest is no doubt positively related to an individual's reasoning ability, it obtains a negative weight so as to suppress the effect of the irrelevant verbal factor in the resultant measure. It is interesting to note that in other areas, suppressor effects have proven especially difficult to cross validate (Adkins, 1947), although it appears that a sampling instability criticism here is inappropriate as the suppression effect was, presumably, evidenced at each of the nine age levels.

Although perhaps generalizable, the use of suppressor variables might still result in difficulties should the resultant measures be used as either dependent variables in some research endeavor, or as diagnostic signs for the practitioner. This latter point is best illustrated through an example taken from our own research. As part of a larger research project we sought to identify those variables which

differentiated learning disabled children from low achieving children. To identify these variables a rather extensive cognitive battery was administered to both a learning disabled and low achieving sample as described earlier. As expected, the mean level of performance of the low achieving sample was significantly better or roughly equivalent to the mean level of performance for the learning disabled sample, except on one measure where the learning disabled children performed significantly better than their low achieving peers (see Table 5). The discrepant result was on the W-J Reasoning factor cluster. Due to the rather strong evidence to the contrary we were, of course, reluctant to interpret this finding as an indicator that learning disabled children possessed greater reasoning skills that their, otherwise superior, low achieving counterparts. This reluctance on our part led us to an investigation of the subtest composition of the Reasoning cluster and finally to an analysis of the mean differences between the two groups on the actual Cognitive Ability subtests.

- - - - - - - - - - - - - - - - - -

Insert Table 5 about here

- - - - - - - - - - - - - - - - - -

Inspection of Table 6 reveals that the two groups performed at about the same level on each of the three positively weighted Reasoning subtests, but that the learning disabled children performed much worse on the negatively weighted Antonyms-Synonyms subtest. Thus, the relative advantage of the learning disabled children on the Reasoning cluster was due almost entirely to the relative disadvantage of these children on the Antonyms-Snyonyms subtest. Apparently, the source of the

present difficulty is the deviant factor structure of the learning disabled children. That is, the suppression of an irrelevant factor is conditional upon a specific factor structure. Should a population follow a different factor structure, the application of the suppression weights will result in invalid measures and perhaps, as here, paradoxical research findings.

- - - - - - - - - - - - - - - - - -

Insert Table 6 about here

- - - - - - - - - - - - - - - - - -

External Validity. The third research question addressed the extent to which the W-J factor clusters evidenced external validity for our learning disabled sample. In order to do so we made use of the correlations between the W-J cluster scores and the other cognitive measures presented in Table 7. Although we might expect attenuation of the correlations due to the extreme nature of the sample, we have no a priori reason to expect differential attenuation of the correlations. Thus our emphasis here is on the relative magnitude rather than the absolute size of these correlations.

- - - - - - - - - - - - - - - - - -

Insert Table 7 about here

- - - - - - - - - - - - - - - - - -

As expected, the Verbal Ability cluster correlated most highly with those external measures with high verbal content: the WISC-R Verbal IQ, along with the Information, Similarities, and Vocabulary subtests. The Verbal Ability cluster also did not correlate very highly with measures not thought to be directly related to verbal ability, as indicated by the lower correlations between the Verbal Ability cluster

and the performance subtests of the WISC-R. The evidence for the external validity for the Verbal Ability cluster, taken in conjunction with our factor analytic results, suggests that verbal ability is a well-defined and valid construct for the interpretation of the cognitive performance of learning disabled students.

In contrast to the Verbal Ability cluster, we find little evidence for the validity of the Perceptual Speed cluster. In fact, the Perceptual Speed cluster evidenced only moderately low correlations with other measures of perception, the BVMGT and the VMI. Although the BVMGT and VMI are primarily power measures, the Perceptual Speed cluster does not appear any more lawful if one looks at its correlation with another direct measure of perceptual speed, the WISC-R Coding subtest ($r = .16$). In fact, the Perceptual Speed cluster had its highest correlations with measures of methematical ability or achievement (WISC-R Arithmetic $r = .49$, PIAT Mathematics $r = .39$, Stanford Mathematical Computation $r = .40$). This rather paradoxical result is best explained by noting that the Perceptual Speed cluster score is determined using the Visual Matching and Spatial Relations subtests relatively weighted three to one respectively. Although the heavily weighted Visual Matching subtest would appear upon first inspection to be a measure of perceptual speed, the task requires that the child select an identical pair of number strings. Thus, on the basis of our factor analytic results, we would predict that since the measure of perceptual speed is so highly saturated with this visual matching task, which requires the processing of numbers, it would correlate most highly with other measures which involve the manipulation of numbers. Thus, we find little evidence for the validity of Woodcock's

perceptual speed factor, if not a little evidence for our alternative factor representation.

For the two factor clusters of Reasoning and Memory, our battery lacks external criteria for validation so we will just briefly highlight our results. The Reasoning cluster shows a rather moderate level of correlation with most other measures, although it appeared to be more highly correlated with the mathematical and verbal measures. It is interesting to note in passing that the Reasoning cluster correlated only moderately with WISC-R Similarities, the WISC-R subtest many would consider most reflective of reasoning. Finally, the Memory cluster appears essentially unrelated to the WISC-R subtests and had as its highest correlates the reading achievement measures on the PIAT.

Scholastic Aptitude Clusters

Development. Although Cognitive Ability subtest performance might be efficiently interpreted in terms of a reduced set of dimensions or factors, these might not be the constructs of primary interest to the educational assessor. Rather, the educational need for the assessment and prediction of achievement might be better fulfilled through the use of measured aptitudes, abilities thought to underlie achievement within a specific domain. As a response to this need, Woodcock and Johnson included measures of scholastic aptitude in their battery. In order to understand the strategy adopted in the development of the Aptitude clusters, it is important to first distinguish between achievement and aptitude. Achievement reflects the knowledge and/or skills a child has acquired; it is a function of that child's ability as well as other noncognitive factors such as motivation and educational opportunity. In

contrast, aptitude reflects the child's ability to achieve. In that the non-cognitive determiners of achievement are more often characteristics of the environment rather than the child, a more accurate assessment of what a child can achieve is obtained through the measurement of that child's aptitude. In a sense then, an aptitude is a predictor of future achievement. Woodcock and Johnson's development of aptitude clusters is consistent with this notion of aptitude as a predictor of achievement.

In contrast to the factorially derived clusters, the Scholastic Aptitude cluster domains were determined a priori, and reflect the standard dimensions of scholastic achievement. Thus we are provided with measures of reading aptitude, mathematics aptitude, written language aptitude, and knowledge aptitude. In determining the subtest composition of the aptitude clusters, Woodcock first excluded those Cognitive Ability subtests having common content with the relevant achievement domain. For example, the Quantitative Concepts subtest is excluded from consideration as a measure of mathematics aptitude because the subtest's items are similar to the criterion of mathematics achievement. The exclusion of subtests sharing a common content with the criterion was motivated by a desire to develop measures that could be used in the assessment of children from special populations, such as the learning disabled. That is, the rather poor achievement of the learning disabled child might not provide a true indication of how much that child could achieve.

As in the development of the factor clusters, subtest composition of the Scholastic Aptitude clusters was determined by selecting the two to four non-excluded Cognitive Ability subtests that best predicted test scores on an external measure of achievement (e.g., W-J Achievement

clusters, PIAT, Kay Math. etc.), on half the developmental sample. The remaining 1,000 test results were then used to develop the relative weighting of the subtests in the computation of cluster scores; as before, weights were determined using the results of a multiple regression analysis at each of the nine age levels.

Although the notion that an aptitude can be used to predict future achievement is easily accepted, the notion that a linear combination of cognitive variables that best predicts achievement would provide a measure of aptitude is dubious at best. In fact, since the aptitude measures were developed through the maximization of the correlation between a linear composite of the Cognitive Ability subtests and measured achievement, would it not be optimal to maximize the developmental criterion by taking as one's measure of aptitude the dependent variable itself, measured achievement?

Subtest Composition and Weighting. With respect to the subtest composition and weighting, the Scholastic Aptitude clusters suffer the same difficulties as the factor clusters. As before, there is a large amount of interdependence among the clusters. To some extent this is a result of actual subtest overlap: the Antonyms-Synonyms subtest is a component of every Scholastic Aptitude cluster; the Analogies, Quantitative Concepts, and Visual Matching subtests are components in two of the four Scholastic Aptitude clusters. The Scholastic Aptitude cluster intercorrelations are presented in Table 8. They are alarmingly high and lead one to question the existence of four separate measures of aptitude as opposed to a general propensity towards achievement. Of course, these high intercorrelations might be characteristic of aptitudes

in general, but if so, then the use of four separate clusters, alluringly named to suggest the existence of specific aptitudes, is apt to mislead the practitioner.

- - - - - - - - - - - - - - - - - -

Insert Table 8 about here

- - - - - - - - - - - - - - - - - -

The use of multiple regression with highly intercorrelated variables has been widely criticized as producing a set of weights with a large amount of sampling variability (Schmidt, 1971; Wainer, 1976). To test the appropriateness of the weighting system used in deriving the Scholastic Aptitude clusters, we sought to replicate the weights found in the developmental sample by comparing this set of weights to the optimal set of weights for our learning disabled sample, and the unit weighting of the subtests. For each aptitude, the dependent variable was the relevant W-J Achievement cluster. The results of this analysis are presented in Table 9.

- - - - - - - - - - - - - - - - - -

Insert Table 9 about here

- - - - - - - - - - - - - - - - - -

The first column presents the cross-validation correlations of the originally developed weighted scheme. In the second column the multiple correlations between the optimal weighting of the subtests in the Aptitude cluster and the Achievement cluster are presented. The third column presents the correlation between the simple linear sum of the Aptitude cluster subtests and the Achievement cluster. Finally, the F ratios used to test the significance between the optimal weighting strategy and the original weights are presented in the fourth column.

There were no significant differences between the correlations obtained using the recommended weights and those obtained when the subtests were optimally weighted. That is, optimal weights could effect no greater prediction of achievement than those weights proposed by Woodcock. Although this finding provides some evidence for the sampling stability of the recommended weighting scheme, there is little evidence for the incremental validity of the differential weights, since the unit weights predicted as well, if not better, than the differential weights.

External Validity. An examination of the relationship between the Scholastic Aptitude clusters and other external measures of achievement leads one to further question the validity of the clusters for a learning disabled sample (see Table 10). When analyzing each Aptitude cluster across the external measures of achievement, a number of issues become apparent. The Mathematics Aptitude Cluster indeed is most highly related to the mathematics achievement on the W-J and the PIAT. However, it is also moderately related to reading achievement as measured by the W-J Reading Achievement cluster and the PIAT Reading Comprehension subtest. While the Reading Aptitude cluster is most highly related to the W-J Reading Achievement cluster, it is no more highly related to reading achievement measured by the PIAT than it is to mathematics achievement measured by the W-J, PIAT, or Stanford. Even greater difficulty exists for the Written Language Aptitude cluster; it is a far better predictor of achievement in mathematics, as evidenced by its relationship with the W-J Mathematics Achievement cluster ($r = .66$), the PIAT Mathematics subtest ($r = .71$), and the Stanford Computation Subtest ($r = .56$), than it is a predictor of achievement in the area of

written language as measured by the W-J Written Language Achievement cluster (r = .48), and the PIAT Spelling subtest (r = .30). Finally, an investigation of the relationships between the Knowledge Aptitude cluster and the external measures of achievement (with the exception of the PIAT Spelling subtest) reveals an equal level of prediction.

- - - - - - - - - - - - - - - - -

Insert Table 10 about here

- - - - - - - - - - - - - - - - -

When looking at the results across external measures, one must view the distinction between these four aptitude measures with suspicion. This is further evidenced by looking at the prediction of specific achievements across clusters. For example, for all measures of mathematics achievement, the Written Language Aptitude cluster is the best predictor. Although the Reading Aptitude cluster is the best predictor of the W-J Reading Achievement cluster, it is not clearly a superior predictor than the other clusters. Furthermore, the best predictor of reading achievement as measured by the PIAT is the Knowledge Aptitude cluster; even the Written Language Aptitude cluster predicts reading achievement better than the Reading Aptitude cluster. Similarly, the Written Language Aptitude cluster is the worst predictor of W-J Written Language Achievement; although it is the best predictor of PIAT Spelling, it predicts only moderately.

Achievement Clusters

Development. At each step, the development of the Achievement clusters was guided by rational choice and expert judgment. The selection of the relevant domains to be measured reflects the standard fare

of the educational system: reading, mathematics, written language, and knowledge. Perhaps the major innovation here is the addition of a fifth cluster, Skills, which is designed to measure those simple tasks which reflect preschool achievement. Subtest composition was based upon expert judgment, selecting those achievement subtests which were relevant to the domain in question. Finally, the subtests composing each cluster were equally weighted in the computation of the achievement cluster scores; this is in contrast to the use of differential weighting schemes for both the factor and aptitude clusters.

Unlike the factor clusters, whose nature must be identified, the achievement domains are provided for by the educational system. That is, the achievement domains represented in the W-J are appropriate for the assessment of any child enrolled in our educational system, learning disabled or not. Alternatively, whereas we would desire to weight the subtests in the factor and aptitude clusters in some optimal fashion, the lack of any available ultimate criteria for achievement precludes the search for optimal weighting schemes. Thus, an equal weighting scheme seems as good, if not more sensible, than any alternative method of subtest weighting.

If there would be any issue related to the development of the Achievement clusters, it would be with respect to subtest composition. It is our position that a determination of the validity of the Achievement cluster subtest composition must be made in terms of a specific educational program. That is, although reading is a relevant achievement domain for all children, the method for assessing reading achievement should not necessarily be the same for all children. In fact,

Jenkins and Pany (1978) demonstrated that pupil performance on alternative reading achievement tests is very much a function of the nature of the reading curriculum in which they were enrolled.

In contrast to the high degree of overlap in the subtest composition of the factor and aptitude clusters, there is little actual subtest overlap in the achievement clusters. The reading, mathematics, written language, and knowledge clusters share no common subtests, although the skills cluster, being comprised of the easiest subtests from the reading, mathematics, and written language clusters, shares one subtest in common with each. Although there is little actual subtest overlap, the intercorrelations between the Achievement clusters were characteristically high in both the present sample and the original standardization group (see Table 11). Again this could be due to the general nature of achievement and should not necessarily be interpreted as a liability of the W-J achievement measures. That this might in fact be the case is illustrated in Table 12, which gives the intercorrelations for the PIAT subtests in our sample; again there is a marked degree of association between these achievement measures.

- - - - - - - - - - - - - - - - - - - -

Insert Table 11 and 12 about here

- - - - - - - - - - - - - - - - - - - -

External Validity. Consistent with the position stated above, any general statements about the validity of the Achievement measures would be inappropriate. Instead, each user should decide whether the measures possess sufficient validity based upon specific content considerations. Despite the specific nature of the validity of the Achievement measures it is interesting to inspect the correlations be-

tween these measures and other widely used achievement measures. These are provided in Table 13. Note that in each case the Achievement cluster having the highest correlation with another measure of achievement was that achievement cluster from the same domain as the criterion.

- - - - - - - - - - - - - - - - - -

Insert Table 13 about here

- - - - - - - - - - - - - - - - - -

Conclusions

The complicated statistical methods used in the Woodcock-Johnson Psycho-Educational Battery to derive cluster measures introduces far greater complexity and confusion than is required. With respect to the Cognitive Factor clusters the following conclusions can be made.

1. There is not a sufficient amount of information reported in the test manual to allow the user to determine the degree to which the Factor Cluster scores measure the originally identified factors.
2. Our failure to replicate the factor structure found in the developmental sample with our learning disabled sample calls into question the appropriateness of the Factor Cluster constructs for the interpretation of the performance of learning disabled children on the 12 Cognitive Ability subtests.
3. The use of suppression weights could lead to paradoxical research findings and invalid measures when used with populations characterized by unique factor structures.

4. Although the Verbal Ability cluster showed the predicted pattern of correlations with external criteria, the Perceptual Speed factor cluster was more highly related to measures not directly associated with the Perceptual Speed cluster than with those measures directly related to it.

With respect to the Scholastic Aptitude clusters the following conclusions can be made.

1. Woodcock and Johnson propose a rather questionable operational definition of an aptitude.
2. The extremely high intercorrelations of the Scholastic Aptitude clusters suggest that, rather than four distinct aptitudes, a general propensity towards achievement is measured.
3. Although the optimal subtest weighting for our learning disabled sample was not significantly superior to the weighting system proposed by Woodcock and Johnson, there was no support for the incremental validity of this weighting system over a simple unit weighting of the subjects.
4. The association between the Aptitude clusters and external measures of achievement provide little evidence for either the differential predictability of achievement or the existence of four distinct and valid measures of aptitude.

Finally, with respect to the achievement clusters, the following

conclusion can be made:

> Although the Achievement clusters evidence a valid pattern of relationships with external criteria and suffer none of the developmental difficulties of the Cognitive Ability factor or Scholastic Aptitude clusters, their use should be based upon the content considerations of a specific educational curriculum rather than on any general statement about their validity.

Although we found little to recommend the use of the W-J cluster scores in our research, we do feel that the test holds some promise for those who would use the battery with special populations. The 22 subtests represent a comprehensive sampling of the cognitive domains and were developed using the results of recent advances in psychometric and latent trait theory. Unfortunately, at present there is little reported psychometric evidence regarding the norms, reliability, and validity of the subtest scores. This difficulty could be easily remedied by re-examining the test results of the carefully selected developmental sample. Finally, for those who are wary of making generalizations from specific cognitive measures, the accumulated empirical evidence regarding the use of differential weights suggests that simple linear composites based upon content considerations might prove superior. After all, one of the greatest achievements in cognitive test development, the Weschler, does not require complex combination rules for the formation of composites. Rather, the simple rationally determined composites employed have proved both valid and interpretable.

References

Adkins, D. C. Construction and analysis of achievement tests. U.S. Government Printing Office, 1947.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multi-trait-multi-method matrix. Psychological Bulletin, 1959, 56, 81-105.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. Manual for Kit of Factor-Referenced Cognitive Test. Princeton: Educational Testing Service, 1976.

Garai, J. E., & Scheinfeld, A. Some differences in mental and behavioral traits. Genetic Psychology Monographs, 1968, 77, 169-299.

Green, B. F. On the factor score controversy. Psychometrika, 1976, 41, 263-266.

Harris, C. W. (Ed.). Problems in measuring change. Madison: University of Wisconsin Press, 1963.

Jenkins, J., & Pany, D. Standardized achievement tests: How useful for special education? Exceptional Children, 1978, 44(6), 448-453.

Schmidt, F. L. The relative efficiency of regression and simple unit predictor weights in applied differential psychology. Educational and Psychological Measurement, 1971, 31, 699-714.

Wainer, H. Estimating coefficients in linear models: It don't make no nevermind. Psychological Bulletin, 1976, 83, 213-217.

Woodcock, R. W. Development and standardization of the Woodcock-Johnson Psycho-Educational Battery. Boston: Teaching Resources, 1978.

Woodcock, R. W., & Johnson, M. B. Woodcock-Johnson Psycho-Educational Battery. Hingham, Mass.: Teaching Resources, 1978.

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1977.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1979.

Footnote

Special appreciation is extended to Don Allen, Pam Ashley, Nicki Cohen, Susan Epps, Nancy Fox, Jean Greener, Kathy Lowry, Nancy McGregor, Steve Poland, Peggy Potter, Paula Savage, and Sherry Schwartz for their assistance in data collection and scoring.

Table 1

Subtests of the W-J Parts One and Two

Part One
Cognitive Battery Subtests

Subtest	Description
1. Picture Vocabulary	Subject names a visually presented stimulus.
2. Spatial Relations	Subject selects a series of component stimuli from a group to form a whole, matching a keyed shape such as a triangle; this subtest is timed.
3. Memory for Sentences	Subject must remember and repeat a string of orally presented words.
4. Visual Auditory Learning	Subject must pair a symbol with familiar words to make sentences.
5. Blending	Subject must combine orally presented syllables and/or phonemes to make a whole word.
6. Quantitative Concepts	Subject must answer questions, presented both orally and visually, related to math; no calculations are required.
7. Visual Matching	Subject must select the pair of identical digit strings from a group of six.
8. Antonyms - Synonyms	Part A: Subject provides an antonym for an orally presented word; Part B: Subject provides a synonym for an orally presented word.
9. Analysis - Synthesis	Subject must use a key to solve visually presented items involving equivalency between colors.
10. Numbers Reversed	Subject must repeat a string of digits in the order opposite that of presentation.
11. Concept Formation	Subject must identify the inclusion rule when given instances and non-instances of the concept.
12. Analogies	Subject must respond with a word that completes an analogy.

Part Two
Achievement Battery Subtests

Subtest	Description
13. Letter-Word Identification	Subject must name visually presented words and letters.
14. Word Attack	Subject must read nonsense words phonetically.
15. Passage Comprehension	Subject must read a short passage and supply a missing word that gives the passage meaning.
16. Calculation	Subject must do math calculations in addition, subtraction, multiplication, and division, as well as some geometric, trigometric, and calculus problems.
17. Applied Problems	Subject must solve practical problems in math.
18. Dictation	Subject must write letters (capital and small), punctuation marks, correct word forms, and spell words.
19. Proofing	Subject must point out errors in written passages, such as incorrect punctuation, word use, and spelling.
20. Science	Subject must answer information items from the biological and physical sciences.
21. Social Studies	Subject must respond to information items from geography, government, economics, etc.
22. Humanities	Subject must answer information items from the domains of art, music, and literature.

Table 2

Cluster Composition of Woodcock-Johnson

		Weight
Cognitive Factor Clusters		
Verbal Ability:		
1.	Picture Vocabulary	.74
8.	Antonyms - Synonyms	.66
9.	Analysis - Synthesis	-.40
Reasoning:		
8.	Antonyms - Synonyms	-.38
9.	Analysis - Synthesis	.44
11.	Concept Formation	.50
12.	Analogies	.44
Perceptual Speed:		
2.	Spatial Relations	.25
7.	Visual Matching	.75
Memory:		
3.	Memory for Sentences	.39
10.	Numbers Reversed	.61
Scholastic Aptitude Clusters		
Reading Aptitude:		
4.	Visual - Auditory Learning	.23
5.	Blending	.17
8.	Antonyms - Synonyms	.42
12.	Analogies	.18
Math Aptitude:		
7.	Visual Matching	.26
8.	Antonyms - Synonyms	.39
9.	Analysis - Synthesis	.27
11.	Concept Formation	.08
Written Language Aptitude:		
6.	Quantitative Concepts	.47
7.	Visual Matching	.15
8.	Antonyms - Synonyms	.29
10.	Numbers Reversed	.08

Knowledge Aptitude:

3.	Memory for Sentences	.10
6.	Quantitative Concepts	.28
8.	Antonyms - Synonyms	.46
12.	Analogies	.16

Achievement Clusters

Reading:

13.	Letter Word Identification	.33
14.	Word Attack	.33
15.	Passage Comprehension	.33

Math:

16.	Calculation	.50
17.	Applied Problems	.50

Written Language:

18.	Dictation	.50
19.	Proofing	.50

Knowledge:

20.	Science	.33
21.	Social Studies	.33
22.	Humanities	.33

Skills:

13.	Letter - Word Identification	.33
17.	Applied Problems	.33
18.	Dictation	.33

Table 3

Varimax Rotated Factor Loadings for the W-J Cognitive Subtests

	Factor 1	Factor 2	Factor 3	Communalities
Picture Vocabulary	.46615	.23625	-.10488	.28
Spatial Relations	.11972	.72579	.00537	.54
Memory for Sentences	.56220	-.06789	.22716	.37
Visual-Auditory	.19650	.40074	.20785	.24
Blending	.40053	.03300	.01942	.16
Quantitative Concepts	.34019	.45818	.53453	.61
Visual Matching	-.14304	.37608	.53033	.44
Antonyms-Synonyms	.70615	.31856	.14674	.62
Analysis-Synthesis	.43517	.54146	.16206	.50
Numbers Reversed	.18044	-.02529	.69773	.52
Concept Formation	.69455	.31891	.07999	.59
Analogies	.62650	.22850	.16044	.47

Table 4

A Comparison of the Intercorrelations of the Factor Clusters in the Learning Disabled Sample with the 5th Grade Developmental Group*

	Verbal Ability	Reasoning	Perceptual Speed	Memory
Verbal Ability	—	.21	.02	.20
Reasoning	.26	—	.17	.27
Perceptual Speed	.20	.31	—	.26
Memory	.37	.42	.37	—

*Values below the diagonal refer to the correlations reported by Woodcock for the fifth graders in his developmental sample (N = 595) while values above the diagonal are for the present sample of learning disabled children (N = 50).

Table 5

Mean Differences on Woodcock-Johnson Factorially Derived Clusters

Scale	Low Achievement Mean (N=49)	Learning Disabled Mean (N=50)	Mean Difference
Verbal Ability	103.3	92.7	4.59
Reasoning	94.7	101.7	-2.25
Perceptual Speed	98.1	93.4	2.08
Memory	101.6	92.2	2.99

Table 6

Mean Differences for the Learning Disabled and the Low Achievement Group for Woodcock-Johnson Cognitive Subtests

	Low Achievement Group (N=49)	Learning Disabled Group (N=50)	Difference t
Picture Vocabulary	18.3	17.2	2.3
Spatial Relations	39.5	38.5	0.97
Memory for Sentences	13.3	11.6	3.50
Visual-Auditory Learning	114.9	114.5	0.23
Blending	16.3	15.7	1.94
Quantitative Concepts	22.0	20.7	2.25
Visual Matching	16.8	16.0	1.8
Antonyms-Synonyms	20.6	17.5	4.13
Analysis-Synthesis	17.4	18.0	-0.87
Numbers Reversed	7.0	6.2	1.81
Concept Formation	16.2	16.1	0.10
Analogies	15.3	15.5	-0.27

Table 7

Correlations of Woodcock-Johnson Factor Clusters with Other Cognitive Tests in a Learning Disabled Sample (N=50)

Woodcock-Johnson Factor Cluster / External Measure	Verbal Ability	Reasoning	Perceptual Speed	Memory
WISC-R Full	.61	.50	.44	.18
WISC-R Verbal	.66	.50	.33	.25
WISC-R Performance	.38	.35	.43	.06
WISC-R Information	.69	.39	.07	.14
WISC-R Similarities	.54	.33	.19	.16
WISC-R Arithmetic	.15	.35	.49	.17
WISC-R Vocabulary	.41	.29	.01	.15
WISC-R Comprehension	.32	.42	.05	.12
WISC-R Picture Completion	.39	.20	-.07	.14
WISC-R Picture Arrangement	.16	.21	.03	-.05
WISC-R Block Design	.24	.24	.20	.04
WISC-R Object Assembly	.15	.15	.19	-.16
WISC-R Coding	-.09	.09	.16	-.12
PIAT Mathematics	.46	.44	.39	.26
PIAT Reading Recognition	.32	.01	.23	.35
PIAT Reading Comprehension	.52	.14	.29	.35
PIAT Spelling	-.07	-.05	.38	.23
PIAT General Information	.68	.24	.14	.00
Stanford Math Computation	.20	.26	.40	.25
Stanford Math Concepts	.13	.52	.26	.03
BENDER (Errors)	-.34	-.35	.07	-.20
BEERY Visual-Motor Integration	.26	.48	.27	.16

Table 8

A Comparison of the Intercorrelations of the Factor Cluster in the Learning Disabled Sample with the 5th Grade Developmental Group*

	Math	Reading	Written Language	Knowledge
Math	-	.79	.78	.84
Reading	.80	-	.62	.86
Written Language	.87	.83	-	.85
Knowledge	.83	.93	.93	-

* Values below the diagonal refer to the correlations reported by Woodcock for the fifth graders in his developmental sample (N=594), while values above the diagonal are for the present sample of learning disabled children (N=50).

Table 9

Cross Validation of the Woodcock-Johnson Aptitude Cluster Weights

Aptitude Cluster	Simple Correlation with Corresponding Woodcock-Johnson Achievement Cluster	Multiple Correlation with Corresponding Woodcock-Johnson Achievement Cluster*	Simple Correlation Using Unit Weights	F**
Reading Aptitude	.55	.61	.58	1.22
Mathematics Aptitude	.57	.62	.61	1.06
Written Language Aptitude	.48	.61	.55	2.48

*Multiple correlation represents the maximum correlation between the achievement cluster and a linear combination of the subtests of the cognitive aptitude cluster.

**All F ratios have 4 and 44 degrees of freedom. To be significant, the observed F would need to exceed $F_{.95}(4.44) = 2.61$.

Table 10

Correlation of W-J Scholastic Aptitude Clusters with Achievement Measures in a Sample of Learning Disabled Children (N=50)*

	MATH				READING			WRITTEN LANGUAGE		KNOWLEDGE
	WJA Math	PIAT Math	Stanford Math Concepts	Stanford Math Computation	WJA Reading	PIAT Reading Recognition	PIAT Reading Comprehension	WJA Written Language	PIAT Spelling	PIAT General Information
W-J Mathematics Aptitude	.57	.61	.34	.40	.46	.25	.42	.51	.07	.53
W-J Reading Aptitude	.46	.44	.27	.42	.55	.33	.42	.54	.07	.42
W-J Written Language Aptitude	.66	.71	.35	.56	.42	.34	.48	.48	.30	.41
W-J Knowledge Aptitude	.59	.63	.28	.46	.52	.40	.53	.59	.16	.60

*Underlined correlations refer to the highest correlation in that column.

Table 11

A Comparison of the Intercorrelations of the Achievement Clusters in the Learning Disabled Sample with the 5th Grade Developmental Group*

	Reading	Mathematics	Written Language	Skills
Reading	--	.36	.74	.79
Mathematics	.60	--	.50	.69
Written Language	.80	.64	--	.75
Skills	.91	.77	.89	--

*Values below the diagonal refer to correlations reported by Woodcock for the fifth graders in his developmental sample (N=590), while values above the diagonal are for the present sample of learning disabled children (N=50).

Table 12

Intercorrelations of the PIAT Achievement Scales

in a Learning Disabled Sample (N=50)

	Mathematics	Reading Recognition	Reading Comprehension	Spelling	General Information
Mathematics	--				
Reading Recognition	.33	--			
Reading Comprehension	.56	.77	--		
Spelling	.24	.62	.49	--	
General Information	.49	.31	.48	.05	--

Table 13

Correlation of the Woodcock-Johnson Achievement Clusters with Other Achievement Measures in a Learning Disabled Sample (N=50)*

W-J Achievement Cluster \ Other Achievement Measures	MATHEMATICS			READING		OTHER	
	PIAT Mathematics	Stanford Mathematical Computations	Stanford Mathematical Concepts	PIAT Reading Recognition	PIAT Reading Comprehension	PIAT Spelling	PIAT General Information
Mathematics	<u>.70</u>	<u>.61</u>	<u>.54</u>	.20	.37	.25	.35
Reading	.38	.35	.35	<u>.77</u>	<u>.73</u>	.46	.38
Written Language	.50	.46	.39	.62	.61	<u>.55</u>	<u>.41</u>
Skills	.62	.50	.48	.69	.69	.51	.37

* Underlined correlations denote the largest column entry.

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $2 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor
Institute for Research on Learning Disabilities
350 Elliott Hall
75 East River Road
University of Minnesota
Minneapolis, Minnesota 55455

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

*Ysseldyke, J. E. Research and development plan: University of Minnesota Institute for Research on Learning Disabilities (Monograph No. 1). June, 1978.

*Ysseldyke, J. E., & Thurlow, M. L. (Eds.). Synthesis of the knowledge base: Identification and assessment of learning disabled children (Monograph No. 2). June, 1978.

*Ysseldyke, J. E. Organizational plan for Minnesota's Institute for Research on Learning Disabilities (Monograph No. 3). June, 1978.

*Ysseldyke, J. E., & Thurlow, M. L. Specific investigations to be completed during years two and three (Monograph No. 4). June, 1978.

*Ysseldyke, J. E., & Thurlow, M. L. Training opportunities at Minnesota's Institute for Research on Learning Disabilities (Monograph No. 5). June, 1978.

*Ysseldyke, J. E. Client base (Monograph No. 6). June, 1978.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

**Thurlow, M., & Ysseldyke, J. Assessment in the child service demonstration centers for learning disabled children (Research Report No. 2). March, 1979.

*As part of its continuation proposal, the Institute was required to prepare these monographs. Because they are part of the proposal, they are not available for general distribution.

**This research report is not being distributed by the Institute. Requests for it should be directed to: The NETWORK, 290 S. Main Street, Andover, Massachusetts 01810.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota round-table conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.